(NASA-CR-146354) MEASURES AND MODELS FOR ANGULAR CORRELATION AND ANGULAR-LINEAR CORRELATION (Wisconsin Univ.) 32 p HC $4.00 CSCL 12A
N76-16884
Unclas
G3/65 13526

Department of Statistics

University of Wisconsin
Madison, Wisconsin

TECHNICAL REPORT NO. 434

January 1976

MEASURES AND MODELS FOR ANGULAR CORRELATION AND ANGULAR-LINEAR CORRELATION

by

Richard A. Johnson and Thomas Wehrly
University of Wisconsin-Madison

* This research was supported by the National Aeronautics Space Administration under Grant NSG-1196.

Abstract

Population models for dependence between two angular measurements and for dependence between an angular and a linear observation are proposed. The method of canonical correlations first leads to new population and sample measures of dependence in this latter situation. An example relating wind direction to the level of a pollutant is given. Next, applied to pairs of angular measurements, the method yields previously proposed sample measures in some special cases and a new sample measure in general.

Key words:

angular correlation

angular-linear correlation

cannonical correlations

distributions on torus

distributions on cylinder

AMS 1970 Subject Classification primary 62H20 62F10 secondary 62G10.

1. Introduction

This investigation centers on the problem of correlation, or dependence, for circular random variables. Although several nonparametric sample measures of dependence have already been proposed for angular observations, there seems to be no literature that treats models for correlation between a circular random variable and a linear random variable. Here we introduce a measure of dependence between circular and linear observations and a similar measure for dependence between two sets of angular observations based on the method of canonical correlation. The asymptotic distribution of the measures is discussed. Some population models are introduced to illustrate the proposed methods. These are among the first population models for random vectors taking values on a cylinder or on a torus and should prove useful in future studies dealing with other population correlation measures.

2. Canonical Correlation Applied to a Random Variable on the Circle and a Random Variable on the Line

An interesting problem is determining a measure of dependence between θ, a random variable taking values on the unit circle, and X, a random variable taking values on the line. We introduce $\underset{\sim}{Y}' = (\cos\theta, \sin\theta)$ to represent θ as a unit vector. We now wish to determine $\underset{\sim}{a}$ such that $\underset{\sim}{a}'\underset{\sim}{Y}$ and X have maximum correlation.

We define the covariance matrix of $(\underset{\sim}{Y}'\ \ X)'$ by

$$\Sigma = \left(\begin{array}{cc|c} \text{var}(\cos\theta) & \text{cov}(\cos\theta, \sin\theta) & \text{cov}(\cos\theta, X) \\ \text{cov}(\cos\theta, \sin\theta) & \text{var}(\sin\theta) & \text{cov}(\sin\theta, X) \\ \hline \text{cov}(\cos\theta, X) & \text{cov}(\sin\theta, X) & \text{var}(X) \end{array}\right) = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{12}' & \sigma^2 \end{array}\right) \quad (1)$$

Let $\underset{\sim}{a}$ be a fixed vector and b be a constant. After imposing the conditions that $\underset{\sim}{a}' \Sigma_{11} \underset{\sim}{a} = 1$, $b^2 \sigma^2 = 1$, the maximum correlation is given by the largest solution of the determinantal equation (c. f. Anderson (1958), Chapter 12)

$$\begin{vmatrix} -\lambda \ \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda \sigma^2 \end{vmatrix} = 0. \qquad (2)$$

Since the correlation $\rho(\underset{\sim}{a}' \underset{\sim}{Y}, X)$ is scale invariant, we can, instead, uniquely determine $\underset{\sim}{a}$ and b by imposing the restrictions that $\underset{\sim}{a}' \underset{\sim}{a} = 1$ and $b^2 = 1$. The maximum correlation subject to these constraints is the same as that obtained by using the usual constraints. We are thus finding the maximum $\rho[\cos(\theta - \alpha), X]$ over all angles α. The angular-linear correlation ρ_{AL} is thus defined by

$$\rho_{AL} = \max_{\alpha} \rho[\cos(\theta - \alpha), X] \qquad (3)$$

In order to make statistical inferences concerning ρ_{AL}, we may either use results related to a specific model such as maximum likelihood estimation theory or else large sample approximations based on the estimated covariance matrix. As is usual with applications of the canonical correlation method as a descriptive measure, one does not need to assume any specific population form but only use the sample covariance matrix given below by (4). The large sample approximations will enable us to determine confidence bounds for ρ_{AL} as well as test for independence.

We consider a random sample (θ_i, X_i) where θ_i is taken from a distribution on the circle and X_i is taken from a distribution on the line. Consider the representation $\underset{\sim}{Y}_i = (\cos \theta_i, \sin \theta_i)'$, $i = 1, \ldots, n$.

Let $\overline{C} = \frac{1}{n}\sum_{i=1}^{n} \cos\theta_i$, $\overline{S} = \frac{1}{n}\sum_{i=1}^{n} \sin\theta_i$, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. The sample covariance matrix for $\underset{\sim}{Z}_i = (\underset{\sim}{Y}_i' \vdots X_i)'$, $i = 1, \ldots, n$, is given by

$$S_n = \left(\begin{array}{cc|c} \frac{1}{n}\Sigma(\cos\theta_i - \overline{C})^2 & \frac{1}{n}\Sigma(\cos\theta_i - \overline{C})(\sin\theta_i - \overline{S}) & \frac{1}{n}\Sigma(\cos\theta_i - \overline{C})(X_i - \overline{X}) \\ \frac{1}{n}\Sigma(\cos\theta_i - \overline{C})(\sin\theta_i - \overline{S}) & \frac{1}{n}\Sigma(\sin\theta_i - \overline{S})^2 & \frac{1}{n}\Sigma(\sin\theta_i - \overline{S})(X_i - \overline{X}) \\ \hline \frac{1}{n}\Sigma(\cos\theta_i - \overline{C})(X_i - \overline{X}) & \frac{1}{n}\Sigma(\sin\theta_i - \overline{S})(X_i - \overline{X}) & \frac{1}{n}\Sigma(X_i - \overline{X})^2 \end{array}\right)$$

$$= \left(\begin{array}{cc|c} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ \hline s_{31} & s_{32} & s_{33} \end{array}\right) = \left(\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right). \tag{4}$$

Then the sample angular-linear correlation r_{AL} is the largest solution to

$$\begin{vmatrix} -\hat{\lambda} S_{11} & S_{12} \\ S_{21} & -\hat{\lambda} S_{22} \end{vmatrix} = 0. \tag{5}$$

We wish to determine the asymptotic distribution of r_{AL}.

Now $\sqrt{n}\; S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\underset{\sim}{Z}_i - \underset{\sim}{\mu})(\underset{\sim}{Z}_i - \underset{\sim}{\mu})' - \sqrt{n}(\overline{\underset{\sim}{Z}}_n - \underset{\sim}{\mu})(\overline{\underset{\sim}{Z}}_n - \underset{\sim}{\mu})'$

where $\overline{\underset{\sim}{Z}}_n = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{Z}_i$, $\underset{\sim}{\mu} = E\underset{\sim}{Z}_i$. The first term of $\sqrt{n}\; S_n$ is asymptotically normal according to the central limit theorem, and the last term goes to zero by the weak law of large numbers. Consequently $\sqrt{n}\,(S_n - \Sigma)$ converges to a normal distribution with mean $\underset{\sim}{0}$. Set $\underset{\sim}{T}_n = (T_1, T_2, \ldots, T_6)'$ equal to $(s_{11}, s_{22}, s_{33}, s_{12}, s_{13}, s_{23})'$

and $\underset{\sim}{\sigma} = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}, \sigma_{23})'$. By looking at the appropriate entries of S and Σ, we see $\sqrt{n}\,(\underset{\sim}{T}_n - \underset{\sim}{\sigma}) \to N(0, B)$ where $B = E(\underset{\sim}{U}\,\underset{\sim}{U}') - \underset{\sim}{\sigma}\,\underset{\sim}{\sigma}'$ and $\underset{\sim}{U} = [(\cos\theta_1 - E\cos\theta_1)^2,\ (\sin\theta_1 - E\sin\theta_1)^2,\ (X_1 - EX_1)^2,$ $(\cos\theta_1 - E\cos\theta_1)(\sin\theta_1 - E\sin\theta_1),\ (\cos\theta_1 - E\cos\theta_1)(X_1 - EX_1),$ $(\sin\theta_1 - E\sin\theta_1)(X_1 - EX_1)\,]'$.

Writing (5) in terms of $\underset{\sim}{T}_n$ and solving for $\hat{\lambda}^2 = r_{AL}^2$, we obtain

$$r_{AL}^2 = \frac{T_1 T_6^2 + T_2 T_5^2 - 2 T_4 T_5 T_6}{T_3 (T_1 T_2 - T_4^2)} . \tag{6}$$

Next we define $g(\underset{\sim}{T}_n) = r_{AL}^2$ where r_{AL}^2 is defined by (6). Then letting $\underset{\sim}{\phi} = \left.\frac{\partial g(\underset{\sim}{T})}{\partial \underset{\sim}{T}}\right|_{\underset{\sim}{T} = \underset{\sim}{\sigma}}$, we have $\sqrt{n}\,(g(\underset{\sim}{T}_n) - \rho_{AL}^2) \xrightarrow{\mathcal{L}} N(0, \underset{\sim}{\phi}' B \underset{\sim}{\phi})$ (c.f. Anderson (1958), p. 76) and by taking the square root transformation, we have $\sqrt{n}\,(\sqrt{g(\underset{\sim}{T}_n)} - \rho_{AL}) \xrightarrow{\mathcal{L}} N(0, \rho_{AL}^{-2}\,\underset{\sim}{\phi}'\,B\,\underset{\sim}{\phi})$. We thus obtain

$$\frac{\sqrt{n}\,(r_{AL} - \rho_{AL})}{v^{\frac{1}{2}}(\underset{\sim}{T}_n)} \xrightarrow{\mathcal{L}} N(0, 1)$$

where $v(\underset{\sim}{T}_n) = v(\Sigma)\big|_{\Sigma = S_n} = \underset{\sim}{\phi}'\,B\,\underset{\sim}{\phi}\,\rho_{AL}^{-2}\big|_{\Sigma = S_n}$ (c.f. C.R. Rao (1973), p 387). Based on the asymptotic normality, a $100(1-\alpha)\%$ confidence interval is given by

$$r_{AL} - \frac{v^{\frac{1}{2}}(\underset{\sim}{T}_n)\, z_{\alpha/2}}{\sqrt{n}} < \rho_{AL} < r_{AL} + \frac{v^{\frac{1}{2}}(\underset{\sim}{T}_n)\, z_{\alpha/2}}{\sqrt{n}}$$

where $z_{\alpha/2}$ is the upper $\frac{\alpha}{2}$ point of a standard normal distribution.

An Example

The wind direction and ozone concentration were observed at a weather station in Milwaukee at 4 day intervals from April 18 to June 29, 1975, at 6 a.m.

Wind direction in degrees	327,	90.9,	88.2,	305,	344,	270,	66.8,	20.5,
Ozone Concentration	28.0,	85.2,	80.5,	4.66,	45.9,	12.7,	72.5,	56.6,
Wind direction in degrees	281,	8.04,	204,	86.4,	335,	18.1,	56.7,	6.03,
Ozone concentration	31.5,	112,	20.0,	72.5,	16.0,	45.9,	32.6,	56.6,
Wind direction in degrees	11.5,	27.0,	84.4.					
Ozone Concentration	52.6,	91.8,	55.2.					

The correlation between X and θ is $r_{AL} = .72$ with a 75% confidence bound $.49 < \rho_{AL} < .96$ and 95% bound $.32 < \rho_{AL} < 1.00$. The angle α of maximum $r(\cos(\theta-\alpha), X) = r_{AL}$ is 74°.

3. A Population Model for Line and the Circle

We now consider (θ, X) having the partly wrapped bivariate normal distribution as described in Appendix 3. We let $\underset{\sim}{Y} = (\cos\theta \; \sin\theta)'$. By using the moments (A.12) and elementary trigonometric identities, the covariance matrix (1) becomes

$$\Sigma = \begin{pmatrix} \frac{1}{2}(1-e^{-\sigma_1^2})(1-e^{-\sigma_1^2}\cos 2\mu_1) & -\frac{1}{2}(1-e^{-\sigma_1^2})e^{-\sigma_1^2}\sin 2\mu_1 & -e^{-\frac{\sigma_1^2}{2}}\rho\sigma_1\sigma_2\sin\mu_1 \\ -\frac{1}{2}(1-e^{-\sigma_1^2})e^{-\sigma_1^2}\sin 2\mu_1 & \frac{1}{2}(1-e^{-\sigma_1^2})(1+e^{-\sigma_1^2}\cos 2\mu_1) & e^{-\frac{\sigma_1^2}{2}}\rho\sigma_1\sigma_2\cos\mu_1 \\ -e^{-\frac{\sigma_1^2}{2}}\rho\sigma_1\sigma_2\sin\mu_1 & e^{-\frac{\sigma_1^2}{2}}\rho\sigma_1\sigma_2\cos\mu_1 & \sigma_2^2 \end{pmatrix}.$$

The determinantal equation (2) reduces to

$$-\frac{\lambda}{2}(1-e^{-\sigma_1^2})\sigma_2^2\left[\frac{\lambda^2}{2}(1-e^{-2\sigma_1^2}) - e^{-\sigma_1^2}\rho\sigma_1^2\right] = 0 .$$

This has roots $\lambda = 0, \pm \dfrac{\rho\sigma_1}{\sqrt{\sinh\sigma_1^2}}$. The maximum root is $\rho_{AL} = \dfrac{|\rho|\,\sigma_1}{\sqrt{\sinh\sigma_1^2}}$.

Consideration of $\rho[X, \cos(\theta-\alpha)]$, where α is a constant, leads to

$$\rho[X, \cos(\theta-\alpha)] = \frac{-\sqrt{2}\,\sin(\mu_1-\alpha)\,\rho\,\sigma_1\,e^{-\frac{\sigma_1^2}{2}}}{\sqrt{(1-e^{-\sigma_1^2})(1-e^{-\sigma_1^2}\cos 2(\mu_1-\alpha))}} .$$

For this to equal ρ_{AL}, we need both $\cos 2(\mu_1-\alpha) = -1$ and $\sin(\mu_1-\alpha) = -\dfrac{|\rho|}{\rho}$. These are satisfied by $\alpha = \mu_1 + \frac{\pi}{2}$ if $\rho > 0$ and $\alpha = \mu_1 - \frac{\pi}{2}$ if $\rho < 0$. We thus obtain $\max_{\alpha} \rho\,[\cos(\theta-\alpha), X] = \begin{cases} \rho\,[\cos(\theta-\mu_1-\frac{\pi}{2}), X] = \rho[\sin(\theta-\mu_1), X] & \text{if } \rho > 0 \\ \rho\,[\cos(\theta-\mu_1+\frac{\pi}{2}), X] = -\rho\,[\sin(\theta-\mu_1), X] & \text{if } \rho < 0 \end{cases}$

The maximum correlation is obtained by centering the θ variable and then rotating by an additional $\pi/2$.

4. Some Previous Measures for Dependence of Angular Observations

The existing literature deals exclusively with nonparametric sample measures of dependence. Epp, Tukey, and Watson (1971) proposed a permutation test for pairs of directional observations. Let $(\underset{\sim}{X}_1, \underset{\sim}{Y}_1), \ldots, (\underset{\sim}{X}_n, \underset{\sim}{Y}_n)$ be a sample of pairs of unit vectors. Then they suggest that a suitable test statistic is $L = \sum_{i=1}^{n} \underset{\sim}{X}_i' \underset{\sim}{Y}_i$, or equivalently, $\sum_{i=1}^{n} \cos\theta_i$ where θ_i is the angle between $\underset{\sim}{X}_i$ and $\underset{\sim}{Y}_i$. Approximations for the permutation distribution are given for both statistics.

Downs, Liebman, and McKay (1967), Downs (1974), and Stephens (1973) develop methods of measuring the rotational correlation. Let $(\underset{\sim}{X}_1, \underset{\sim}{Y}_1), \ldots, (\underset{\sim}{X}_n, \underset{\sim}{Y}_n)$ be a sample of pairs of unit vectors. It is desired to determine the extent to which each $\underset{\sim}{X}_i$ is a constant rotation of $\underset{\sim}{Y}_i$. That is, whether each $\underset{\sim}{Y}_i = H\underset{\sim}{X}_i$ where H is an orthogonal matrix with determinant one. To do this, they find the orthogonal matrix $\hat{H}$ which minimizes

$$f = \sum_{i=1}^{n} (\underset{\sim}{Y}_i - H\underset{\sim}{X}_i)'\,(\underset{\sim}{Y}_i - H\underset{\sim}{X}_i)$$

or, equivalently, that $\hat{H}$ which maximizes

$$r^* = \sum_{i=1}^{n} (H\underset{\sim}{X}_i)'\,\underset{\sim}{Y}_i \qquad (7)$$

where the minimum and maximum are taken over all orthogonal matrices H.

Both Downs (1967) and Stephens (1973) give a method for obtaining H.

The first authors propose a rotational correlation coefficient which is analogous to the usual sample correlation coefficient and is defined by

$$r = \frac{\sum_{i=1}^{n} (\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}})' \hat{H} (\underset{\sim}{X}_i - \hat{\underset{\sim}{X}}) \qquad |\hat{H}|}{\left[\sum_{i=1}^{n} (\underset{\sim}{X}_i - \hat{\underset{\sim}{X}})' (\underset{\sim}{X}_i - \hat{\underset{\sim}{X}}) \sum_{i=1}^{n} (\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}})' (\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}) \right]^{\frac{1}{2}}} \tag{8}$$

where $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ are unit vectors having the same direction as resultant vectors of the sets $\{\underset{\sim}{X}_i\}$ and $\{\underset{\sim}{Y}_i\}$, respectively. Stephens proposes the measures $r = \sum_{i=1}^{n} (\hat{H} \underset{\sim}{X}_i)' \underset{\sim}{Y}_i$ where $\hat{H}$ is the orthogonal matrix which maximizes (7) and $r_+ = \sum_{i=1}^{n} (H^* \underset{\sim}{X}_i)' \underset{\sim}{Y}_i$ where H^* is the orthogonal matrix with determinant one which maximizes (7).

Rothman (1971) adapts a test for independence based on the empirical cumulative distribution function to a test of coordinate independence for a sample on the torus. Mardia (1975) defines a correlation coefficient for circular data based on the ranks of the observations. Rao and Puri (197?) propose a test for coordinate independence based on the number of observations falling in half-circles. They derive the asymptotic distribution of the test statistic and a computational form of this test in terms of the X and Y spacings .

5. Canonical Correlations Applied to Bivariate Circular Data.

Let θ_1 and θ_2 be two random variables which take values on the unit circle. One may consider the representation

$$\underset{\sim}{X}_1 = \begin{pmatrix} \cos\theta_1 \\ \sin\theta_1 \end{pmatrix}, \qquad \underset{\sim}{X}_2 = \begin{pmatrix} \cos\theta_2 \\ \sin\theta_2 \end{pmatrix} . \tag{9}$$

This provides a 1-1 correspondence between each angle in the interval $[0, 2\pi)$ and the set of unit vectors. Based on the representation, the objective of the cannonical correlation method is to find $\underset{\sim}{a}$ and $\underset{\sim}{b}$ such that $\underset{\sim}{a}' \underset{\sim}{X}_1$ and $\underset{\sim}{b}' \underset{\sim}{X}_2$ have maximum correlation.

We let $\underset{\sim}{X}' = (\underset{\sim}{X}_1' \vdots \underset{\sim}{X}_2')$ and denote the covariance matrix of $\underset{\sim}{X}$ by

$$\Sigma = \left(\begin{array}{cc|cc} \mathrm{Var}(\cos\theta_1) & \mathrm{Cov}(\cos\theta_1, \sin\theta_1) & \mathrm{Cov}(\cos\theta_1, \cos\theta_2) & \mathrm{Cov}(\cos\theta_1, \sin\theta_2) \\ \mathrm{Cov}(\sin\theta_1, \cos\theta_1) & \mathrm{Var}(\sin\theta_1) & \mathrm{Cov}(\sin\theta_1, \cos\theta_2) & \mathrm{Cov}(\sin\theta_1, \sin\theta_2) \\ \hline \mathrm{Cov}(\cos\theta_2, \cos\theta_1) & \mathrm{Cov}(\cos\theta_2, \sin\theta_1) & \mathrm{Var}(\cos\theta_2) & \mathrm{Cov}(\cos\theta_2, \sin\theta_2) \\ \mathrm{Cov}(\sin\theta_2, \cos\theta_1) & \mathrm{Cov}(\sin\theta_2, \sin\theta_1) & \mathrm{Cov}(\sin\theta_2, \cos\theta_2) & \mathrm{Var}(\sin\theta_2) \end{array}\right)$$

$$= \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{21} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right) . \qquad (10)$$

Further, let $\underset{\sim}{a}$ and $\underset{\sim}{b}$ be constant vectors. We wish to find $\underset{\sim}{a}$ and $\underset{\sim}{b}$ which maximize the correlation between $\underset{\sim}{a}' \underset{\sim}{X}_1$ and $\underset{\sim}{b}' \underset{\sim}{X}_2$. Generally the maximizing combination is made unique by imposing the conditions that $\underset{\sim}{a}' \Sigma_{11} \underset{\sim}{a} = 1$ and $\underset{\sim}{b}' \Sigma_{22} \underset{\sim}{b} = 1$. The maximum correlation is then given by the largest root of the determinantal equation,

$$\begin{vmatrix} -\lambda\,\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\,\Sigma_{22} \end{vmatrix} = 0 . \qquad (11)$$

Since the correlation coefficient is scale invariant, we can instead impose the restrictions on $\underset{\sim}{a}$ and $\underset{\sim}{b}$ that $\underset{\sim}{a}' \underset{\sim}{a} = \underset{\sim}{b}' \underset{\sim}{b} = 1$. In this case,

$\underset{\sim}{a}$ and $\underset{\sim}{b}$ can be represented by

$$\underset{\sim}{a} = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \qquad \underset{\sim}{b} = \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}. \tag{12}$$

Moreover, $\underset{\sim}{a}'\underset{\sim}{X}_1 = \cos(\theta_1 - \alpha)$, $\underset{\sim}{b}'\underset{\sim}{X}_2 = \cos(\theta_2 - \beta)$, and the problem becomes that of maximizing the correlation between $\cos(\theta_1 - \alpha)$ and $\cos(\theta_2 - \beta)$ over all $\alpha, \beta \in [o, 2\pi)$. Consequently, we define the angular canonical correlation between θ_1 and θ_2 by

$$\rho_A = \sup_{\alpha, \beta \in [o, 2\pi)} \rho[\cos(\theta_1 - \alpha), \cos(\theta_2 - \beta)]. \tag{13}$$

This measure is obviously invariant under rotations of θ_1 and θ_2.

Unlike the sample measures (7) and (8) discussed above, it is, in general, necessary to rotate both coordinates in order to obtain the maximum correlation. The necessity of the two rotations will be shown for the bivariate wrapped normal model of Section 7.2 and for the example of section 6. In section 7.1, one rotation will be seen to be sufficient to obtain maximum correlation for certain models with uniform marginals.

One important practical feature of using the method of canonical correlation to measure angular correlation is that it enables one to find the angular correlation by using standard statistical programs for canonical correlations. The angles α and β which give the maximum correlation can easily be found from such programs by converting the coefficients for $\underset{\sim}{a}$ and $\underset{\sim}{b}$ to unit vectors and then finding the α and β which satisfy (12).

We also consider the following heuristic justification for the above method for determining angular correlation. Suppose one wishes to find the functions $f(\theta_1)$ and $g(\theta_2)$ which have the greatest correlation.

The Fourier expansions of $f(\theta_1)$ and $g(\theta_2)$ have the form

$$f(\theta_1) - \sum_{n=o}^{\infty} (a_n \cos n\theta_1 + b_n \sin n\theta_1), \quad g(\theta_2) = \sum_{n=o}^{\infty} (c_n \cos n\theta_2 + d_n \sin n\theta_2).$$

Ignoring second order and higher terms, we obtain

$f(\theta_1) \cong a_o + a_1 \cos \theta_1 + b_1 \sin \theta_1 = a_o + A_1 \cos(\theta_1 - \alpha)$ and

$g(\theta_2) \cong c_o + c_1 \cos \theta_2 + d_1 \sin \theta_2 = c_o + A_2 \cos(\theta_2 - \beta)$. Then

$$\max_{f, g} \rho[f(\theta_1), g(\theta_2)] \cong \max_{\substack{a_o, A_1, \alpha, \\ c_o, A_2, \beta}} \rho[a_o + A_1 \cos(\theta_1 - \alpha),\ c_o + A_2 \cos(\theta_2 - \beta)]$$

$$= \max_{\alpha, \beta} \rho\,[\cos(\theta_1 - \alpha),\ \cos(\theta_2 - \beta)] = \rho_A.$$

We note that our proposed measure of correlation can be considered as a measure of rotational correlation. The following lemma shows that perfect correlation occurs if and only if one angle is a constant rotation of the other angle.

<u>Lemma 5.1</u> If θ_1 and θ_2 are circular random variables whose distributions have support $[o, 2\pi)$, then $\rho_A = 1 \Leftrightarrow \theta_2 = \theta_1 + \delta \pmod{2\pi}$, with probability one, for some δ.

<u>Proof</u> First assume that $\rho_A = 1$. Then there are some α, β such that $\rho[\cos(\theta_1 - \alpha), \cos(\theta_2 - \beta)] = 1$, and by the Cauchy - Schwartz inequality, $\cos(\theta_2 - \beta) = A \cos(\theta_1 - \alpha) + B$ with probability one for some constants A and B. Since the support includes (θ_1, θ_2) with θ_1 and θ_2 both taking all values in $[o, 2\pi)$, the requirement that $\cos(\theta_2 - \beta)$ vary from -1 to $+1$ leads to $B = 0$, $A = \pm 1$. For $A = 1$, $\cos(\theta_2 - \beta) = \cos(\theta_1 - \alpha)$ and $\theta_2 - \beta = \theta_1 - \alpha \pmod{2\pi}$. We take $\delta = \beta - \alpha \pmod{2\pi}$. For $A = -1$, $\cos(\theta_2 - \beta) = -\cos(\theta_1 - \alpha) = \cos(\theta_1 - \alpha - \pi)$, and we take $\delta = \beta - \alpha - \pi \pmod{2\pi}$.

If $\theta_2 = \theta_1 + \delta \pmod{2\pi}$ with probability one, then

$$1 \geq \rho_A = \max_{\alpha,\beta} \rho[\cos(\theta_1 - \alpha), \cos(\theta_2 - \beta)] = \max_{\alpha,\beta} \rho[\cos(\theta_1 - \alpha), \cos(\theta_1 + \delta - \beta)]$$

$$\geq \rho[\cos\theta_1, \cos\theta_1] = 1.$$

Remark

We also note that a linear relationship between θ_1 and θ_2 does not imply perfect correlation. In fact we can have $\theta_2 = 2\theta_1 \pmod{2\pi}$ and $\rho_A = 0$. Let θ_1 have the uniform distribution on $[0, 2\pi)$ and $\theta_2 = 2\theta_1 \pmod{2\pi}$. Then it is easy to check that $\text{cov}[\cos(\theta_1 - \alpha), \cos(2\theta_1 - \beta)] = 0$ for all α, β and hence, $\rho_A = 0$. However, independence clearly implies $\rho_A = 0$.

6. Inference from Sample Angular Correlations

In this section we outline a method of obtaining the large sample distribution of the angular correlation coefficient, r_A. This will provide an asymptotic method of finding confidence intervals for r_A when the underlying distribution is unknown. In the case where the family of underlying distributions is known, some statistic based on these distributions, such as the maximum likelihood estimate of ρ_A, should be used.

We consider a random sample (θ_i, η_i), $i = 1, \ldots, n$, from a bivariate circular distribution. Consider the representation

$$\underset{\sim}{U}_i = \begin{pmatrix} \cos\theta_i \\ \sin\theta_i \end{pmatrix}, \quad \underset{\sim}{V}_i = \begin{pmatrix} \cos\eta_i \\ \sin\eta_i \end{pmatrix}, \quad \underset{\sim}{Z}_i = \begin{pmatrix} \underset{\sim}{U}_i \\ \underset{\sim}{V}_i \end{pmatrix}, \quad i = 1, \ldots, n. \tag{14}$$

Let $\bar{\underset{\sim}{Z}}_n = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{Z}_i$ be the sample mean vector. Then the sample covariance matrix is

$$S_n = \frac{1}{n} \sum_{i=1}^{n} (\underset{\sim}{Z}_i - \overline{\underset{\sim}{Z}}_n)(\underset{\sim}{Z}_i - \overline{\underset{\sim}{Z}}_n)' = \left(\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right) = \left(\begin{array}{cc|cc} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ \hline s_{31} & s_{32} & s_{33} & s_{34} \\ s_{41} & s_{42} & s_{43} & s_{44} \end{array}\right) \qquad (15)$$

We now define

$$\underset{\sim}{T}_n = (s_{11}, s_{22}, s_{33}, s_{44}, s_{12}, s_{13}, s_{14}, s_{23}, s_{24}, s_{34})'$$

The determinantal equation (11) with S_{ij} replacing Σ_{ij} becomes

$$C_1 \hat{\lambda}^4 + C_2 \hat{\lambda}^2 + C_3 = 0 \qquad (16)$$

where C_1, C_2, C_3 can be viewed as functions of $\underset{\sim}{T}_n$. By expanding (11) we obtain

$$C_1 = (T_1 T_2 - T_5^2)(T_3 T_4 - T_{10}^2) \qquad (17)$$

$$C_2 = -T_1 T_4 T_8^2 + 2T_1 T_8 T_9 T_{10} - T_1 T_3 T_9^2 - 2T_5 T_7 T_8 T_{10} - 2T_5 T_6 T_9 T_{10}$$

$$+ 2T_3 T_5 T_7 T_9 + 2T_4 T_5 T_6 T_8 - T_2 T_4 T_6^2 + 2T_2 T_6 T_7 T_{10} - T_2 T_3 T_7^2$$

$$C_3 = (T_6 T_9 - T_7 T_8)^2$$

The solution to (16) becomes $\lambda^2 = (-C_2 \pm \sqrt{C_2^2 - 4C_1 C_3}) / 2C_1$.

Thus, $r_A = \hat{\lambda}$, the largest root of (11), is given by

$$g(\underset{\sim}{T}_n) = \hat{\lambda} = \sqrt{\frac{-C_2 + \sqrt{C_2^2 - 4C_1C_3}}{2C_1}} \quad . \tag{18}$$

Using the same proof as in section 2 above, $\sqrt{n}\,(S_n - \Sigma)$ is asymptotically normal with mean zero and the asymptotic covariance of the $(j,k)^{th}$ entry and the $(\ell, m)^{th}$ entry given by

$$E[(z_j-\mu_j)(z_k-\mu_k)(z_\ell-\mu_\ell)(z_m-\mu_m)] - E[(z_j-\mu_j)(z_k-\mu_k)] \quad E[(z_\ell-\mu_\ell)(z_m-\mu_m)]$$

where

$$z_1 = \cos\theta_i, \;\; z_2 = \sin\theta_i, \;\; z_3 = \cos\eta_i, \;\; z_4 = \sin\eta_i, \;\; \mu_j = E z_j \; .$$

Set $\underset{\sim}{\sigma} = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{44}, \sigma_{12}, \sigma_{13}, \sigma_{14}, \sigma_{23}, \sigma_{24}, \sigma_{34})'$. Then $\sqrt{n}\;(\underset{\sim}{T}_n - \underset{\sim}{\sigma}) \xrightarrow{\mathcal{L}} N(0, B)$ where B is a 10×10 matrix with entries corresponding to those given for the asymptotic distribution of S. Then $\sqrt{n}\;(r_A - \rho_A) \xrightarrow{\mathcal{L}} N(0, \underset{\sim}{\phi}'\, B\, \underset{\sim}{\phi})$ where

$$\underset{\sim}{\phi} = \left.\frac{\partial g(\underset{\sim}{T}_n)}{\partial \underset{\sim}{T}_n}\right|_{\underset{\sim}{T}_n = \underset{\sim}{\sigma}} \quad . \tag{19}$$

One can straightforwardly determine $\underset{\sim}{\phi}$ by using (17), (18), and (19), and thus obtain an asymptotic variance. To find confidence bounds, we let $v(\Sigma) = \underset{\sim}{\phi}'\, B\, \underset{\sim}{\phi}$.

Then $\dfrac{\sqrt{n}\,(\hat{\rho}_A - \rho_A)}{v^{\frac{1}{2}}(S)} \xrightarrow{\mathcal{L}} N(0,1)$ where $v(S) = v(\Sigma)\big|_{\Sigma=S}$. The above method was illustrated in more detail in section 2 dealing with the correlation between X and θ.

Example

The wind direction at 6 a.m. and 12 noon were measured each day at a weather station in Milwaukee for 21 consecutive days. We wish to determine whether the two wind directions are correlated.

Wind Direction in Degrees

6 a.m.	356,	97.2,	211,	232,	343,	292,	157,	302,	335,	302,	324,	84.6,
noon	119,	162,	221,	259,	270,	28.8,	97.2,	292,	39.6,	313,	94.2,	45,
6 a.m.	324,	340,	157,	238,	254,	146,	232,	122,	329			
noon	47,	108,	221,	270,	119,	248,	270,	45,	23.4			

The direction of the resultants are 286° at 6 a.m. and 33° at noon. We obtain $r_A = .5673$, with $.23 < \rho_A < .90$ as a 95% confidence bound for ρ_A. The angles α and β for maximum correlation are $\alpha = 26^\circ$ and $\beta = 58^\circ$.

Suppose instead we used a single rotation procedure as suggested by Downs et al. (1967). Then, if we fix θ_2 (noon readings), the maximum correlation between $\cos(\theta_1 - \alpha)$ and $\cos\theta_2$ is .4479 for $\alpha = -12^\circ$. This is substantially smaller than r_A.

7. Canonical Correlation Applied to Some Bivariate Models on the Torus

Population models of bivariate circular random variables having dependence are indispensible for studying the various measures of correlation. To partially fill a noticeable void in the literature, we introduce the following models.

7.1 Models with Uniform Marginals

In this section we will discuss the application of canonical correlation to models with uniform marginals. Two possible models are given, the second more general than the preceding model.

Suppose (θ_1, θ_2) follows a bivariate circular distribution whose density is given by

$$f(\theta_1, \theta_2) = \frac{1}{4\pi I_0(\kappa)} e^{\kappa \cos(\theta_1 - \theta_2 - \mu_0)}, \quad 0 \le \theta_1, \theta_2 < 2\pi, \tag{20}$$

where $\kappa > 0$ and $0 \leq \mu_o < 2\pi$ are the parameters. This model can be obtained by finding the distribution on the circle which maximizes the entropy, $-\int_0^{2\pi}\int_0^{2\pi} f(\theta_1, \theta_2) \log f(\theta_1, \theta_2)\, d\theta_1 d\theta_2$, subject to the conditions

$$E[\cos(\theta_1-\theta_2)] = A \cos \mu_o , \qquad E[\sin(\theta_1-\theta_2)] = A \sin \mu_o,$$

where A and μ_o are preassigned constants (c. f. Kagan, et al. (1973), p. 409). In this model, the marginal distributions of θ_1 and θ_2 are both uniform while the difference of the angles, $\theta_1-\theta_2$, has a von Mises distribution.

For this model, we now consider the representation given by (9), and let Σ be defined by (10). By straightforward integration, we obtain

$$\Sigma = \left(\begin{array}{cc|cc} \frac{1}{2} & 0 & \frac{A(\kappa)}{2}\cos\mu_o & -\frac{A(\kappa)}{2}\sin\mu_o \\ 0 & \frac{1}{2} & \frac{A(\kappa)}{2}\sin\mu_o & \frac{A(\kappa)}{2}\cos\mu_o \\ \hline \frac{A(\kappa)}{2}\cos\mu_o & \frac{A(\kappa)}{2}\sin\mu_o & \frac{1}{2} & 0 \\ -\frac{A(\kappa)}{2}\sin\mu_o & \frac{A(\kappa)}{2}\cos\mu_o & 0 & \frac{1}{2} \end{array}\right)$$

where $A(\kappa) = I_1(\kappa)/I_o(\kappa)$ and $I_p(\kappa)$ is the modified Bessel function of the first kind and of p^{th} order. The determinantal equation (11) becomes

$$(\lambda^2 - A(\kappa)^2)^2 = 0 .$$

Hence, the maximum correlation between $\underset{\sim}{a}'\underset{\sim}{X}_1$ and $\underset{\sim}{b}'\underset{\sim}{X}_2$ subject to $\underset{\sim}{a}'\underset{\sim}{a} = \underset{\sim}{b}'\underset{\sim}{b} = 1$ is given by $A(\kappa)$. Using the representation (12), the population first canonical correlation is given by $\rho_A = \max_{\alpha,\beta} \rho(\cos(\theta_1-\alpha), \cos(\theta_2-\beta)) = A(\kappa)$.

We now determine the rotations α and β for which the maximum is attained. First of all, $\mathrm{Cov}[\cos(\theta_1-\alpha), \cos(\theta_2-\beta)] = \frac{A(\kappa)}{2}\cos(\alpha-\beta-\mu_o)$, and $\mathrm{Var}[\cos(\theta_1-\alpha)] = \mathrm{Var}[\cos(\theta_2-\beta)] = \frac{1}{2}$. Thus, $\rho[\cos(\theta_1-\alpha), \cos(\theta_2-\beta)] = A(\kappa)\cos(\alpha-\beta-\mu_o)$ which is maximized for $\cos(\alpha-\beta-\mu_o) = 1$ or $\alpha-\beta-\mu_o = 0 \pmod{2\pi}$. Hence, $\alpha-\beta = \mu_o \pmod{2\pi}$ maximizes ρ. Although α, β are not unique mod 2π, one can be chosen arbitrarily and then the other will be fixed. This shows, in addition, that $\max_{\alpha} \rho[\cos(\theta_1-\alpha), \cos\theta_2] = \max_{\alpha,\beta} \rho[\cos(\theta_1-\alpha), \cos(\theta_2-\beta)] = \max_{\beta} \rho[\cos\theta_1, \cos(\theta_2-\beta)]$ for this model. We thus need only rotate one of the two angles to find the maximum correlation.

Next, consider estimating $A(\kappa)$ using a random sample $(\theta_{1,i}, \theta_{2,i})$, $i = 1, \ldots, n$, from this distribution. The maximum likelihood estimate of $A(\kappa)$ (c.f. Mardia, p. 122) is given by $\widehat{A(\kappa)} = \bar{R}$ where $\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$, $\bar{C} = \frac{1}{n}\sum_{i=1}^{n} \cos(\theta_{1,i} - \theta_{2,i})$, and $\bar{S} = \frac{1}{n}\sum_{i=1}^{n} \sin(\theta_{1,i} - \theta_{2,i})$. It is easy to

show that $\bar{R} = \frac{1}{n}\sum_{i=1}^{n} \cos(\theta_{1,i} - \theta_{2,i} - \bar{x}_o)$ where $\bar{x}_o$ is the solution of $\bar{C} = \bar{R}\cos\bar{x}_o$, $\bar{S} = \bar{R}\sin\bar{x}_o$. That is, to estimate $A(\kappa)$, we rotate $\theta_1-\theta_2$ by $\bar{x}_o$, the maximum likelihood estimate of μ_o, and find $\bar{C}^* = \frac{1}{n}\sum_{i=1}^{n} (\cos(\theta_{1,i} - \theta_{2,i} - \bar{x}_o)$. The use of $\bar{R}$ for testing for independence is discussed and tables given by Mardia (1972), page 136 and page 300.

The above method coincides exactly with that suggested by Stevens (1973) who proposed the sample measure $r_+ = \max_{H \text{ rotations}} \sum_{i=1}^{n} (H\underset{\sim}{X}_i)'\underset{\sim}{Y}_i = \max_{H} \sum_{i=1}^{n} \cos\theta_i$ where θ_i is the angle between $H\underset{\sim}{X}_i$ and $\underset{\sim}{Y}_i$ when $\underset{\sim}{X}_i$, $\underset{\sim}{Y}_i$ are unit vectors. To see this, we note that multiplication of $\underset{\sim}{X}_i$ by a rotation matrix H corresponds to subtracting a constant α from $\theta_{1,i}$ where $\underset{\sim}{X}_i = (\cos\theta_{1,i} \sin\theta_{1,i})'$, and

$$\max_{H \text{ rotation}} \sum_{i=1}^{n} (H\underset{\sim}{X}_i)'\underset{\sim}{Y}_i = \max_{\alpha} \sum_{i=1}^{n} \cos(\theta_{1,i} - \theta_{2,i} - \alpha).$$ Differentiating

and setting equal to zero gives the maximum at $\alpha = \bar{x}_o$, so $\frac{1}{n} r_+ = \hat{A}(\kappa)$ for this model.

Next we consider a more general model, still having uniform marginals, with density of the form

$$g(\theta_1, \theta_2) = \frac{1}{2\pi} h(\theta_1 - \theta_2 - \mu_o), \quad 0 \leq \theta_1, \theta_2 < 2\pi,$$

where μ_o is a parameter and $h(\cdot)$ is a circular density; i.e., $h(x) \geq 0$ and $\int_0^{2\pi} h(x)\,dx = 1$. If we let $h(x) = (2\pi I_o(\kappa))^{-1} \exp[\kappa \cos x]$, $0 \leq x < 2\pi$, this model reduces to (20).

Let $A = \int_0^{2\pi} \cos x \; h(x)\,dx$, $B = \int_0^{2\pi} \sin x \; h(x)\,dx$, and define $\underset{\sim}{X}$ by (9) and Σ by (10). The covariance matrix Σ becomes

$$\left(\begin{array}{cc|cc}
\frac{1}{2} & 0 & \frac{1}{2}(A\cos\mu - B\sin\mu) & -\frac{1}{2}(A\sin\mu + B\cos\mu) \\
0 & \frac{1}{2} & \frac{1}{2}(A\sin\mu + B\cos\mu) & \frac{1}{2}(A\cos\mu - B\sin\mu) \\
\hline
 & & \frac{1}{2} & 0 \\
 & \bullet & & \\
 & & 0 & \frac{1}{2}
\end{array}\right)$$

And equation (11), for the canonical correlations, reduces to $(\lambda^2 - (A^2 + B^2))^2 = 0$. Hence, the maximum correlation for $\underset{\sim}{a}'\underset{\sim}{X}_1$, $\underset{\sim}{b}'\underset{\sim}{X}_2$ subject to $\underset{\sim}{a}' \Sigma_{11} \underset{\sim}{a} = \underset{\sim}{b}' \Sigma_{22} \underset{\sim}{b} = 1$, or equivalently subject to $\underset{\sim}{a}'\underset{\sim}{a} = \underset{\sim}{b}'\underset{\sim}{b} = 1$ is $\rho_A = \sqrt{A^2 + B^2}$. Next represent $\underset{\sim}{a}$ and $\underset{\sim}{b}$ by (12). Then

$\rho(\underset{\sim}{a}'\underset{\sim}{X}_1, \underset{\sim}{b}'\underset{\sim}{X}_2) = \rho[\cos(\theta_1 - \alpha), \cos(\theta_2 - \beta)] = A\cos(\alpha - \beta - \mu) + B\sin(\alpha - \beta - \mu)$.
Comparing this with $\sqrt{A^2 + B^2}$, we obtain $\tan(\alpha - \beta - \mu) = \frac{B}{A}$. Again a single rotation of either θ_1 or θ_2 with the other assuming an arbitrary value is sufficient to obtain the maximum correlation.

Remark It seems that the single rotation definition is closely tied to uniform marginals.

7.2 The Wrapped Bivariate Normal Distribution

Let (θ_1, θ_2) be a random vector with the wrapped bivariate normal distribution. (See the appendix for a description of this distribution.) Let $\mathfrak{X}_1$, $\mathfrak{X}_2$, and Σ be defined by (9) and (10) respectively. By using the trigonometric moments (A.10) and some elementary trigonometric identities, we obtain the covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} ,$$

$$\Sigma_{11} = \begin{pmatrix} \frac{1}{2} a\,(1-e^{-\sigma_1^2} \cos 2\mu_1) & -\frac{1}{2} a\, e^{-\sigma_1^2} \sin 2\mu_1 \\ -\frac{1}{2} a\, e^{-\sigma_1^2} \sin 2\mu_1 & \frac{1}{2} a\,(1+e^{-\sigma_1^2} \cos 2\mu_1) \end{pmatrix}$$

$$\Sigma_{12} = \begin{pmatrix} \frac{1}{2} e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}(b \cos(\mu_1+\mu_2)+c \cos(\mu_1-\mu_2)) & \frac{1}{2} e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}(b \sin(\mu_1+\mu_2)-c \sin(\mu_1-\mu_2)) \\ \frac{1}{2} e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}(b \sin(\mu_1+\mu_2)+c \sin(\mu_1-\mu_2)) & \frac{1}{2} e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}(-b \cos(\mu_1+\mu_2)+c \cos(\mu_1-\mu_2)) \end{pmatrix}$$

$$\Sigma_{22} = \begin{pmatrix} \frac{1}{2} d(1-e^{-\sigma_2^2} \cos 2\mu_2) & -\frac{1}{2} d\, e^{-\sigma_2^2} \sin 2\mu_2 \\ -\frac{1}{2} d\, e^{-\sigma_2^2} \sin 2\mu_2 & \frac{1}{2} d\,(1+e^{-\sigma_2^2} \cos 2\mu_2) \end{pmatrix}$$

where $a = 1-e^{-\sigma_1^2}$, $b = e^{-\rho\sigma_1\sigma_2}-1$, $c = e^{\rho\sigma_1\sigma_2}-1$, $d = 1-e^{-\sigma_2^2}$. (21)

After some algebra, the equation given by (11) reduces to

$$\left[\frac{\lambda^2}{4}(1-e^{-\sigma_1^2})^2(1-e^{-\sigma_2^2})^2 - e^{-(\sigma_1^2+\sigma_2^2)}(\cosh \rho\sigma_1\sigma_2-1)^2\right] \times$$

$$\left[\frac{\lambda^2}{4}(1-e^{-2\sigma_1^2})(1-e^{-2\sigma_2^2}) - e^{-(\sigma_1^2+\sigma_2^2)} \sinh \rho\sigma_1\sigma_2\right] = 0$$

whose solutions are

$$\lambda_1 = \frac{\cosh \rho\sigma_1\sigma_2-1}{2 \sinh \frac{\sigma_1^2}{2} \sinh \frac{\sigma_2^2}{2}}, \qquad \lambda_2 = \frac{|\sinh \rho\sigma_1\sigma_2|}{\sqrt{\sinh \sigma_1^2 \sinh \sigma_2^2}}.$$

The correlation λ_1 corresponds to rotating θ_1 by μ_1 and θ_2 by μ_2, that is rotating both variables until they are centered at their means. The correlation λ_2 corresponds to rotating θ_1 by θ_1 by $\mu_1 + \frac{\pi}{2}$ and θ_2 by $\mu_2 + \frac{\pi}{2}$ if ρ is positive and to rotating θ_1 by $\mu_1 - \frac{\pi}{2}$ and θ_2 by $\mu_2 - \frac{\pi}{2}$ if ρ is negative. Thus, $\lambda_1 = \rho[\cos(\theta_1-\mu_1), \cos(\theta_2-\mu_2)]$ and, if $\rho \neq 0$,

$\lambda_2 = \rho[\cos(\theta_1-\mu_1-\frac{|\rho|}{\rho}\ \frac{\pi}{2}), \cos(\theta_2-\mu_2-\frac{|\rho|}{\rho}\ \frac{\pi}{2})]$.

We now give a proposition concerning these correlations and discuss some of their properties.

<u>Proposition 7.1</u>

If $\sigma_1 = \sigma_2$, then $\lambda_2 > \lambda_1$ if $\rho \neq 1$; $\lambda_2 = \lambda_1$ if $\rho = 1$ or if $\rho = 0$.

<u>Proof</u> Assume $\rho \neq 0$.

$$\frac{\lambda_1^2}{\lambda_2^2} = \frac{(\cosh \rho\sigma^2-1)^2}{\sinh^2 \rho\sigma^2} \; \frac{\sinh^2 \sigma^2}{4\sinh^4 \frac{\sigma^2}{2}} = \frac{(\cosh \rho\sigma^2-1)^2}{1-\cosh^2 \rho\sigma^2} \; \frac{\left(e^{\sigma^2} - e^{-\sigma^2}\right)^2}{\left(e^{\frac{\sigma^2}{2}} - e^{-\frac{\sigma^2}{2}}\right)^4}$$

$$= \frac{\cosh \rho\sigma^2-1}{\cosh \rho\sigma^2+1} \; \frac{\left(e^{2\sigma^2}-1\right)^2}{\left(e^{\sigma^2}-1\right)^4} = \frac{\cosh \rho\sigma^2-1}{\cosh \rho\sigma^2+1} \; \frac{\left(e^{2\sigma^2}+2e^{\sigma^2}+1\right)}{\left(e^{2\sigma^2}-2e^{\sigma^2}+1\right)}$$

$$= \frac{(\cosh \rho\sigma^2-1)}{(\cosh \rho\sigma^2+1)} \; \frac{(\cosh \sigma^2+1)}{(\cosh \sigma^2-1)} = \frac{\cosh \sigma^2 \cosh \rho\sigma^2-1+(\cosh \rho\sigma^2-\cosh \sigma^2)}{\cosh \sigma^2 \cosh \rho\sigma^2-1+(\cosh \sigma^2-\cosh \rho\sigma^2)}$$

If $\rho < 1$, then $\cosh \rho\sigma^2 < \cosh \sigma^2$ and $\frac{\lambda_1^2}{\lambda_2^2} < 1$; if $\rho = 1$,

$\cosh \rho\sigma^2 = \cosh \sigma^2$ and $\frac{\lambda_1^2}{\lambda_2^2} = 1$. If $\rho = 0$, $\lambda_1 = \lambda_2 = 0$.

Proposition 7.1 states that if the variances are equal for both marginal distributions, then $\rho_A = \lambda_2$. Moreover, a numerical study indicates that $\lambda_2 > \lambda_1$ for all ρ when σ_1 and σ_2 are relatively close in value, and $\lambda_1 > \lambda_2$ when one variance is very much larger than the other. In other words, in the usual cases where the marginal variances not too different,

$\rho_A = \lambda_2 = \rho[\cos(\theta_1-\mu_1-\frac{\pi}{2}), \cos(\theta_2-\mu_2-\frac{\pi}{2})] = \rho[\sin(\theta_1-\mu_1), \sin(\theta_2-\mu_2)]$.

It can be shown that $\lambda_1 \leq 1$ and $\lambda_2 \leq 1$ with $\lambda_1 = \lambda_2 = 1$ if and only if $\rho = 1$ and $\sigma_1 = \sigma_2$. Thus, $\cos(\theta_1-\alpha)$ and $\cos(\theta_2-\beta)$ can be perfectly correlated only if the underlying bivariate normal distribution is perfectly correlated and has equal variances. In this case, $\theta_1 = \theta_2 + \delta$ for some δ, which agrees with an earlier lemma.

For general α and β, the correlation between $\cos(\theta_1-\alpha)$ and $\cos(\theta_2-\beta)$ is given by

$$\rho[\cos(\theta_1-\alpha), \cos(\theta_2-\beta)] = \frac{e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)}[b\cos(\mu_1-\alpha+\mu_2-\beta) + c\cos(\mu_1-\alpha-(\mu_2-\beta))]}{\sqrt{a(1-e^{-\sigma_1^2}\cos 2(\mu_1-\alpha))\ \ d(1-e^{-\sigma_2^2}\cos 2(\mu_2-\beta))}}$$

where a, b, c, d are defined by (21). Consequently, ρ_A depends on both α and β, and the two rotations are necessary to maximize the correlation.

Acknowledgement We wish to express our appreciation to G. K. Bhattacharyya for helpful and stimulating discussion concerning this work.

Appendix

A.1. Wrapping on the Torus

A method of forming a circular distribution from a distribution on the line is to wrap the distribution on the unit circle. That is, if X is a random variable with c.d.f. $F(x)$ and c.f. $\phi(t)$, then a circular variable is determined from $\theta = X \pmod{2\pi}$. With this transformation, θ has c.d.f.

$$F_\theta(\theta) = \sum_{k=-\infty}^{\infty} (F(\theta + 2\pi k) - F(2\pi k)), \quad 0 \leq \theta < 2\pi,$$

and characteristic function $\phi_p = \phi(p)$ (c.f. Mardia (1972), p. 53).

Analogously, one can form a bivariate circular distribution by wrapping a bivariate distribution, defined on the plane, on the unit torus. Let (X_1, X_2) be a bivariate random vector from a distribution having c.d.f. $F(x_1, x_2)$, density $f(x_1, x_2)$, and characteristic function $\phi(t_1, t_2)$. Let

$$\theta_1 = X_1 \pmod{2\pi}, \quad \theta_2 = X_2 \pmod{2\pi}. \tag{A.1}$$

Then θ_1 and θ_2 are random variables on the circle, and (θ_1, θ_2) is a random vector on the torus. The c.d.f. of (θ_1, θ_2) is

$$F_{\underset{\sim}{\theta}}(\theta_1, \theta_2) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} [F(\theta_1 + 2\pi j, \theta_2 + 2\pi k) - F(2\pi j, \theta_2 + 2\pi k) - F(\theta_1 + 2\pi j, 2\pi k) + F(2\pi j, 2\pi k)].$$

$0 \leq \theta_1, \theta_2 < 2\pi$; and the density is given by

$$f_{\underset{\sim}{\theta}}(\theta_1, \theta_2) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(\theta_1 + 2\pi j, \theta_2 + 2\pi k), \quad 0 \leq \theta_1, \theta_2 < 2\pi.$$

Lemma A 1 The c. f. of (θ_1, θ_2) is $\phi_{p,q} = \phi(p, q)$ where p and q are integers.

Proof

$$\phi_{p,q} = E(e^{ip\theta_1 + iq\theta_2}) = \int_0^{2\pi}\int_0^{2\pi} e^{ip\theta_1 + iq\theta_2} f_{\underset{\sim}{\theta}}(\theta_1, \theta_2)\, d\theta_1 d\theta_2$$

$$= \int_0^{2\pi}\int_0^{2\pi} e^{ip\theta_1 + iq\theta_2} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(\theta_1 + 2\pi j,\ \theta_2 + 2\pi k)\, d\theta_1\, d\theta_2$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \int_0^{2\pi}\int_0^{2\pi} e^{ip\theta + iq\theta_2} f(\theta_1 + 2\pi j,\ \theta_2 + 2\pi k)\, d\theta_1\, d\theta_2$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \int_{2\pi j}^{2\pi(j+1)}\int_{2\pi k}^{2\pi(k+1)} e^{ip\theta_1 + iq\theta_2} f(\theta_1, \theta_2)\, d\theta_1\, d\theta_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{ip\theta_1 + iq\theta_2} f(\theta_1, \theta_2)\, d\theta_1\, d\theta_2 = \phi(p, q).$$

Theorem A2.

$\phi_{p,q}$ determined at integer values of p and q is sufficient to determine the distribution of $\underset{\sim}{\theta}$.

Proof

Define the family of functions

$$u_\rho(\xi, \eta) = \frac{1}{4\pi^2} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \phi_{p,q}\, \rho^{|p| + |q|} e^{-ip\xi - iq\eta}, \quad 0 \leq \rho < 1. \qquad (A.2)$$

If $\phi_{p,q} = 1$, (A. 2) reduces to

$$\frac{1}{4\pi^2} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \rho^{|p| + |q|} e^{-ip\xi - iq\eta} = C(\eta;\rho)\, C(\xi;\rho), \quad 0 \leq \rho < 1. \qquad (A. 3)$$

where $C(\eta;\rho) = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \rho^{|p|} e^{-ip\eta}$ is the p. d. f. of the wrapped Cauchy distribution (or is the Poisson kernel). Substituting $\int_0^{2\pi}\int_0^{2\pi} e^{ip\theta_1+iq\theta_2} dF(\theta_1, \theta_2) = \phi_{p,q}$ into (A. 2), we obtain

$$u_\rho(\xi, \eta) = \frac{1}{4\pi^2} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \int_0^{2\pi}\int_0^{2\pi} e^{i(p\theta_1+q\theta_2)} dF(\theta_1, \theta_2)\, \rho^{|p|+|q|} e^{-ip\xi-iq\eta}$$

$$= \int_0^{2\pi}\int_0^{2\pi} \left(\frac{1}{4\pi^2} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \rho^{|p|+|q|} e^{-ip(\xi-\theta_1)-iq(\eta-\theta_2)} \right) dF(\theta_1, \theta_2)$$

$$= \int_0^{2\pi}\int_0^{2\pi} C(\xi-\theta_1;\rho)\, C(\eta-\theta_2;\rho)\, dF(\theta_1, \theta_2) . \qquad (A. 4)$$

The second equality follows by the dominated convergence theorem, and the last equality holds by applying (A 3).

Using (A. 4) and applying Fubini's Theorem, we obtain

$$\lim_{\rho\to 1} \int_0^{\alpha_1}\int_0^{\alpha_2} u_\rho(\xi, \eta)\, d\xi\, d\eta = \lim_{\rho\to 1} \int_0^{\alpha_1}\int_0^{\alpha_2}\int_0^{2\pi}\int_0^{2\pi} C(\xi-\theta_1;\rho) C(\eta-\theta_2;\rho)\, dF(\theta_1, \theta_2)\, d\xi\, d\eta$$

$$= \lim_{\rho\to 1} \int_0^{2\pi}\int_0^{2\pi}\int_0^{\alpha_1} C(\xi-\theta_1;\rho) d\xi \int_0^{\alpha_2} C(\eta-\theta_2;\rho)\, d\eta \quad dF(\theta_1, \theta_2) . \qquad (A.5)$$

A lemma concerning the Poisson kernel (c. f. Feller (1971), p. 627) states that

$$\lim_{\rho\to 1} \int_0^{\alpha} C(\xi-\theta;\rho)\, d\xi = I_{(0,\alpha]}(\theta) . \qquad (A. 6)$$

The dominated convergence theorem and (A. 6) give

$$\lim_{\rho \to 1} \int_0^{\alpha_1} \int_0^{\alpha_2} u_\rho(\xi, \eta) \; d\xi \, d\eta = \int_0^{2\pi} \int_0^{2\pi} I_{(0, \alpha_1]}(\theta_1) \, I_{(0, \alpha_2]}(\theta_2) \, dF(\theta_1, \theta_2)$$

$$= F(\alpha_1, \alpha_2).$$

Hence, $F(\alpha_1, \alpha_2)$ can be obtained by a limiting process from the $u_\rho(\xi, \eta)$ which depend only on $\phi_{p,q}$.

A. 2. The Wrapped Bivariate Normal Distribution

Let (X_1, X_2) be bivariate normal with mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \qquad \underset{\sim}{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} . \tag{A. 7}$$

Define $\underset{\sim}{\theta}$ by (A. 1) so, according to Lemma A 1, the characteristic function of $\underset{\sim}{\theta}$ is

$$\phi_{p,q} = \exp\{-\tfrac{1}{2}(p^2\sigma_1^2 + 2pq\,\rho\sigma_1\sigma_2 + q^2\sigma_2^2) + i(p\mu_1 + q\mu_2)\} . \tag{A. 8}$$

Since

$$\phi_{p,q} = E(e^{ip\theta_1 + iq\theta_2}) = E\cos(p\theta_1 + q\theta_2) + i\,E\sin(p\theta_1 + q\theta_2) \tag{A. 9}$$

for any random vector $\underset{\sim}{\theta}$, equating (A. 9) and (A. 8), we obtain

$$\begin{aligned} E\cos(p\theta_1+q\theta_2) &= \cos(p\mu_1+q\mu_2)\,\exp\{-\tfrac{1}{2}(p^2\sigma_1^2 + 2pq\,\rho\sigma_1\sigma_2 + q^2\sigma_2^2)\} \\ E\sin(p\theta_1+q\theta_2) &= \sin(p\mu_1+q\mu_2)\,\exp\{-\tfrac{1}{2}(p^2\sigma_1^2 + 2pq\,\rho\sigma_1\sigma_2 + q^2\sigma_2^2)\} \end{aligned} \tag{A. 10}$$

A. 3 A Model for (θ, X)

One model for (θ, X) can be formed from a bivariate normal distribution. Let (Y_1, Y_2) have mean vector $\underset{\sim}{\mu}$ and covariance matrix Σ defined by (A. 7). Let $\theta = Y_1 (\text{mod } 2\pi)$, $X = Y_2$. Then (θ, X) has a distribution with the desired support and having characteristic function

$$\phi(p,t) = \exp\{-\tfrac{1}{2}(p^2\sigma_1^2 + 2pt\rho\sigma_1\sigma_2 + q^2\sigma_2^2) + i(p\mu_1 + t\mu_2)\} . \qquad \text{(A. 11)}$$

The proofs that this is the desired c. f. and that it is necessary to determine $\phi(p,t)$ only for integer p and real t are entirely analogous to those presented in section A. 1. The moments $E(X\cos\theta)$ and $E(X\sin\theta)$ can be easily determined from the c. f.

<u>Lemma A 3</u> If $E|X|$ exists and is finite, then

$$\left.\frac{\partial\phi(p,t)}{\partial t}\right|_{\substack{t=0\\p=1}} = -E(X\sin\theta) + i\,E(X\cos\theta).$$

<u>Proof</u> The c. f. is $\phi(p,t) = E(e^{ip\theta + itX})$. Then

$$\frac{\partial\phi}{\partial t} = \frac{\partial}{\partial t}\int_0^{2\pi}\int_{-\infty}^{\infty} e^{ip\theta + itx}\,dF(\theta, x)$$

$$= \int_0^{2\pi}\int_{-\infty}^{\infty} \frac{\partial}{\partial t} e^{ip\theta + itx}\,dF(\theta, x)$$

$$= \int_0^{2\pi}\int_{-\infty}^{\infty} ix\,e^{ip\theta + itx}\,dF(\theta, x).$$

The interchange of differentiation and integration is justified by the fact $E|X| < \infty$ and the dominated convergence theorem.

If (θ, X) has c.f. given by (A.11),

$$\left.\frac{\partial\phi(p,t)}{\partial t}\right|_{\substack{t=0\\p=1}} = e^{-\frac{\sigma_1^2}{2}} [(-\rho\sigma_1\sigma_2 \cos\mu_1 - \mu_2 \sin\mu_1) + i(\mu_2\cos\mu_1 - \rho\sigma_1\sigma_2 \sin\mu_1)] .$$

Hence, the moments are given by :

$$E(X\cos\theta) = e^{-\frac{\sigma_1^2}{2}} (\mu_2\cos\mu_1 - \rho\sigma_1\sigma_2\sin\mu_1)$$

$$E(X\sin\theta) = e^{-\frac{\sigma_1^2}{2}} (\mu_2\sin\mu_1 + \rho\sigma_1\sigma_2\cos\mu_1) . \qquad \text{(A. 12)}$$

We note that the marginals of this distribution are the wrapped normal with parameters μ_1 and σ_1^2 and the normal distribution with parameters μ_2 and σ_2^2 for θ and X respectively. This can be seen by considering $t = 0$ and $p = 0$ in the c.f. $\phi(p,t)$.

References

1. Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons.

2. Downs, T. D. (1974). Rotational angular correlations, in Biorhythms and Human Reproduction, (M. Ferin et al. Eds.), pp. 97-104. New York: John Wiley and Sons.

3. Downs, T. Liebman, J., and McKay, W. (1967). Statistical methods for vectorcardiogram orientations. Proceedings XIth International Vectorcardiography Symposium. New York: North Holland Publishing Company.

4. Epp, R. J., Tukey, J. W., and Watson, G. S. (1971). Testing unit vectors for correlation. Journal of Geophysical Research, 76, 8480-8483.

5. Feller, W. (1971). An Introduction to Probability Theory and Its Applications, Volume II. New York: John Wiley and Sons.

6. Gould, A. L., (1969). A regression technique for angular variates. Biometrics. 25, 683-700.

7. Kagan, A. M., Linnik, Yu. V. and Rao, C. R. (1973). Characterization Problems in Mathematical Statistics. New York: John Wiley and Sons.

8. Mardia, K. V. (1972). Statistics of Directional Data. New York: Academic Press.

9. Mardia, K. V. (1975). Statistics of direction data (with Discussion), Journal of the Royal Statistical Society B, 37, 349-393.

10. Rao, C. R. (1973). Linear Statistical Inference and Its Applications, 2nd ed. New York: John Wiley and Sons.

11. Rao, J. S. and Puri, M. L., (1975). Problems of association for bivariate circular data and a new test of independence. Technical Report No. 426, Department of Statistics, University of Wisconsin.

12. Rothman, E. D., (1971). Tests of coordinate independence for a bivariate sample on a torus. The Annals of Mathematical Statistics, 42, 1962-1969.

13. Stephens, M. A., (1973). Vector correlation, part 1. Technical Report No. 7, George Washington University. Department of Statistics.

14. Watson, G. S., and Beran, R. J. (1967). Testing a sequence of unit vectors for serial correlation, Journal of Geophysical Research, 72, 5655-9.